



gRaphics

Data-driven story telling

Chloe **Fouilloux** (Part I) and Sara **Hocevar** (Part II)



@ChloeFouilloux



chloefouilloux.com



sara.hocevar@icloud.com

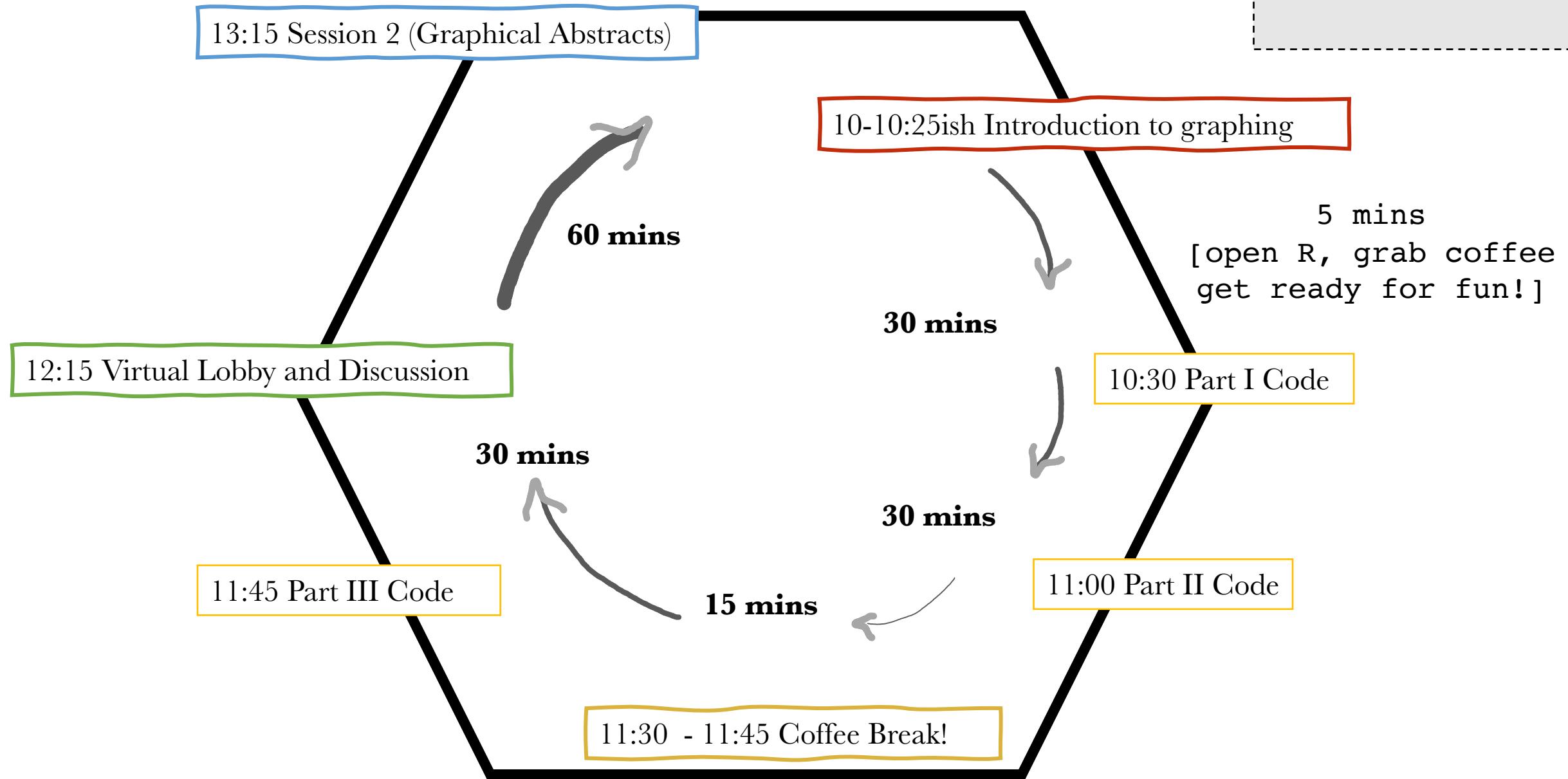


UNIVERSITY OF JYVÄSKYLÄ



Oikos Finland

Session I Schedule



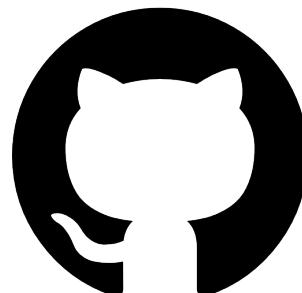
Workshop Accessibility



Live Zoom. Live lecture and Q&A, code troubleshooting, and networking opportunities.



Lecture components later available on YouTube.
Any personal information (chat question, user id) will be removed before upload. (i.e. only my voice and screen will be put in final product).



Accessible code! All code can be freely accessed on GitHub ([chloefouilloux/Oikos_Workshop](https://github.com/chloefouilloux/Oikos_Workshop)).

THE DOMINANT TREE SPECIES' PLANTED ALONG SAN FRANCISCO'S ROADS

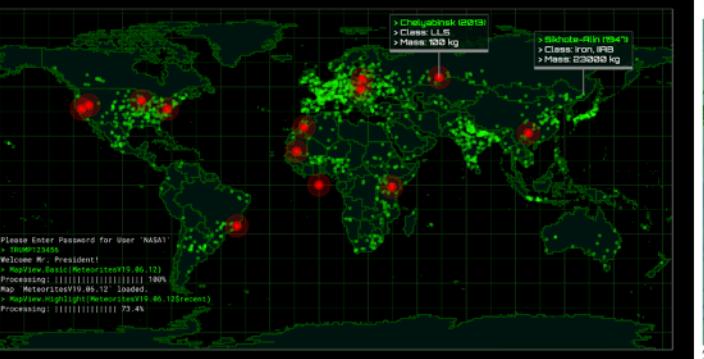
Spatial Data

Categorical data

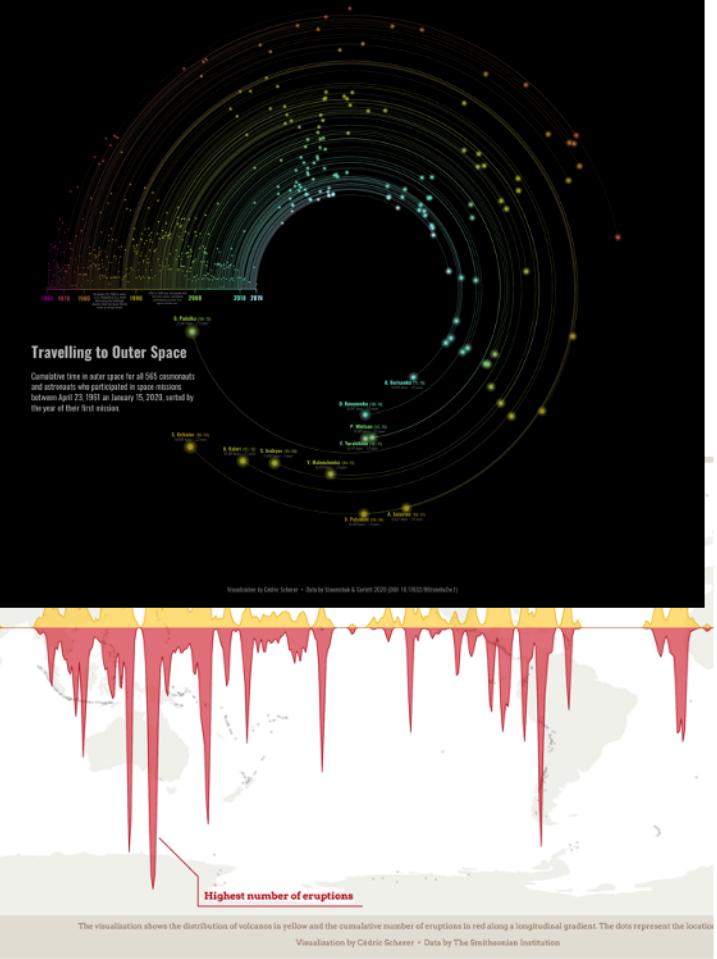
Continuous data



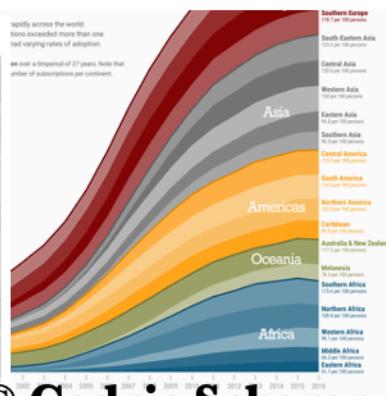
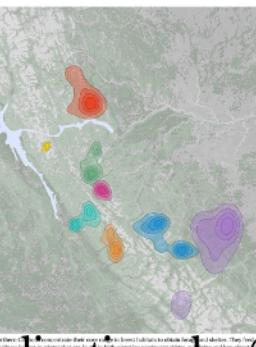
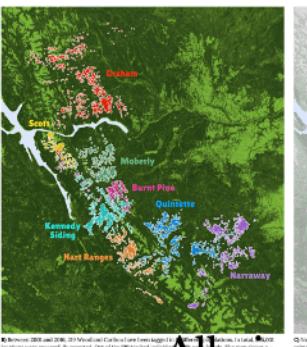
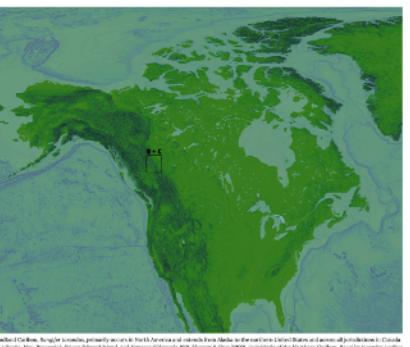
SPACE OBSERVER 3000 X V3.5.1



Provided by Cédric Scherer & National Aeronautics and Space Administration (NASA)

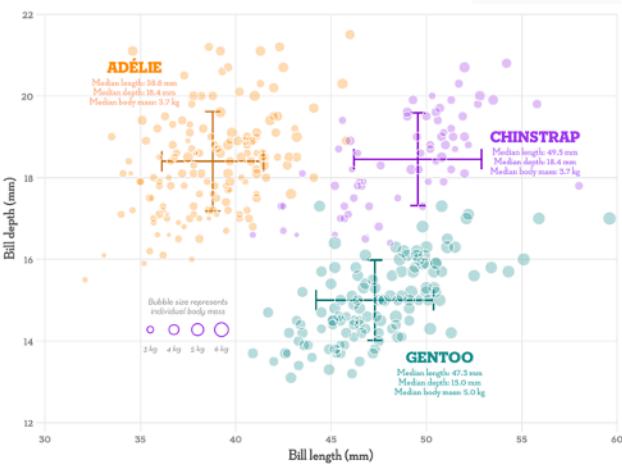


Spatial patterns of South Peace Northern Caribou, *Rangifer tarandus caribou* (pop. 15) in British Columbia 2001–2016

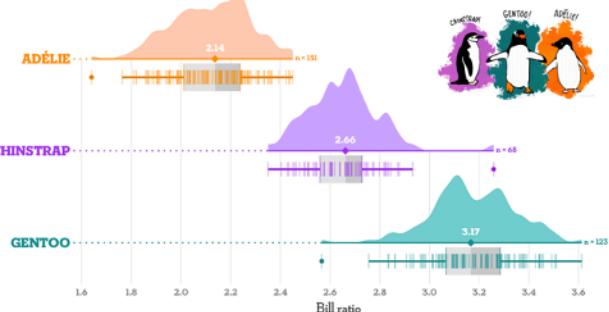


BILL DIMENSIONS OF BRUSH-TAILED PENGUINS
Pygoscelis adeliae (Adélie penguin) • *P. antarctica* (Chinstrap penguin) • *P. papua* (Gentoo penguin)

A. Scatterplot of bill length versus bill depth (error bars show median +/- sd)



B. Distribution of the bill ratio, estimated as bill length divided by bill depth

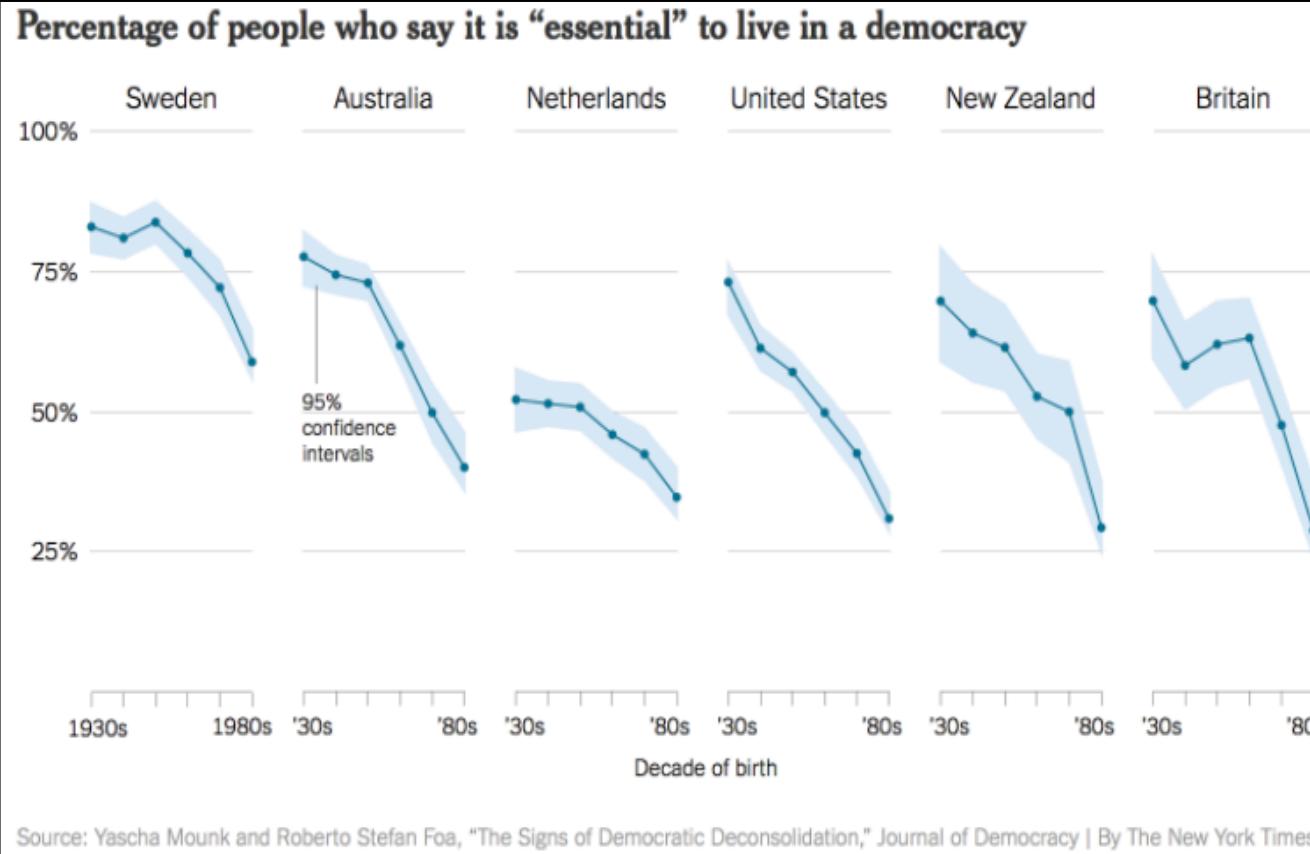


Note: In the original data, bill dimensions are recorded as "culmen length" and "culmen depth". The culmen is the dorsal (upper) ridge of a bird's bill.

Visualization: Cédric Scherer • Data: German, Williams & Fraser (2014) DOI: 10.1371/journal.pone.0090081 Illustrations: Alison Horst

All visualizations by © Cédric Scherer

“How Stable Are Democracies? ‘Warning Signs Are Flashing Red’”(Taub, 2016)



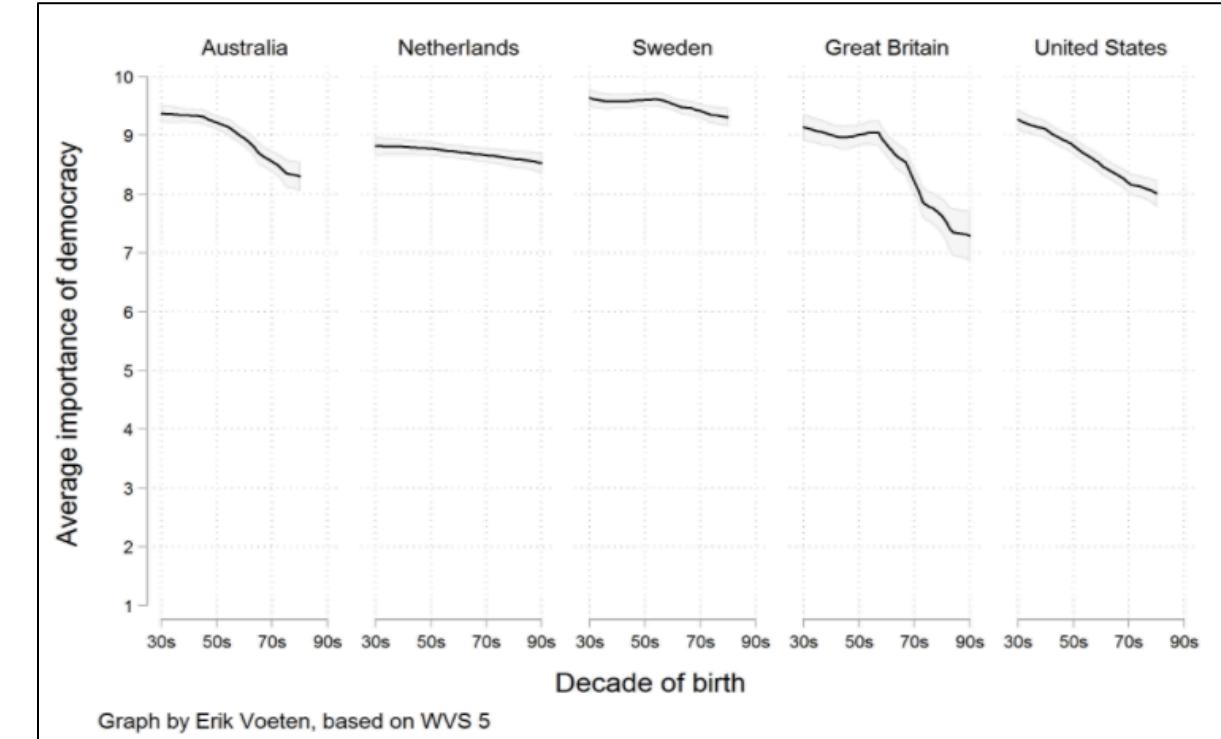
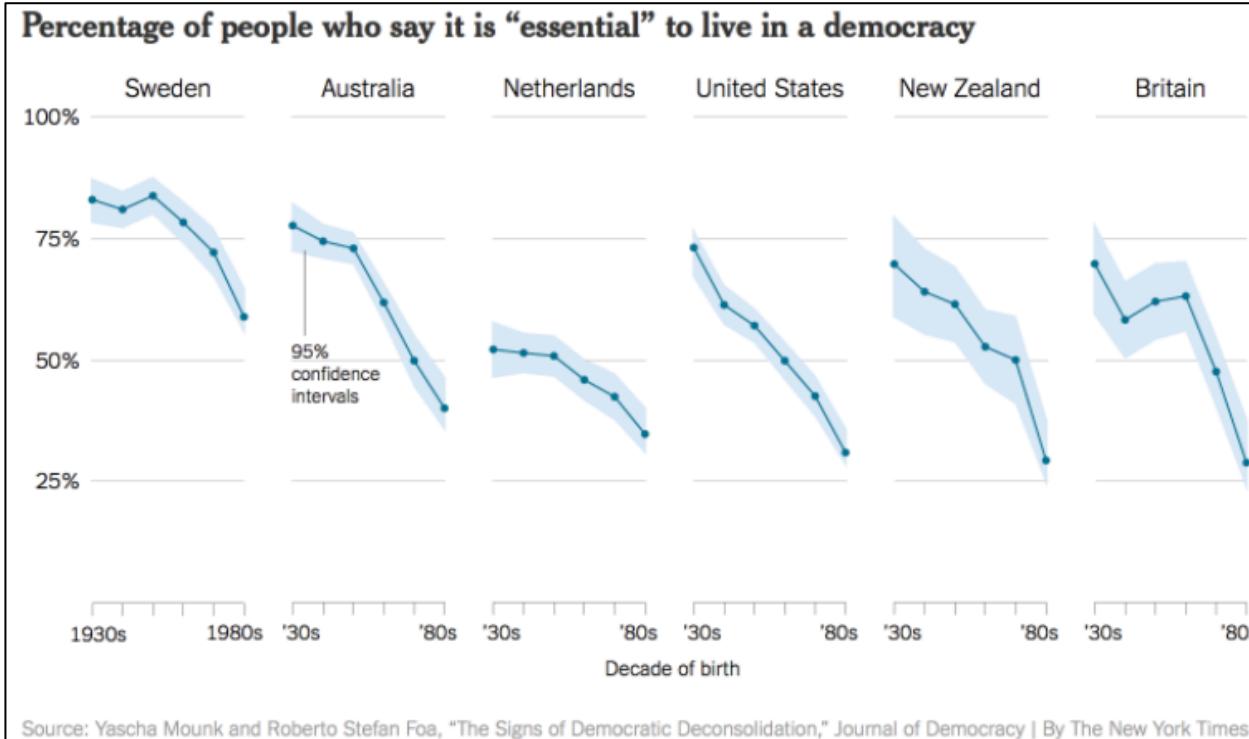
Initial impression is alarming!

Well labeled, shows CIs. . . what could possibly be wrong with this graph?

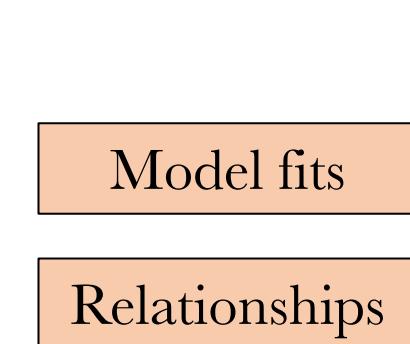
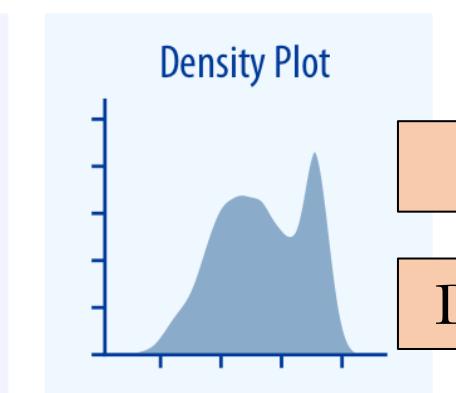
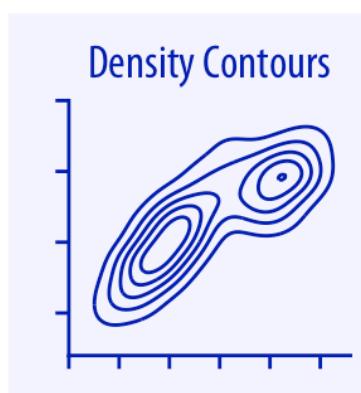
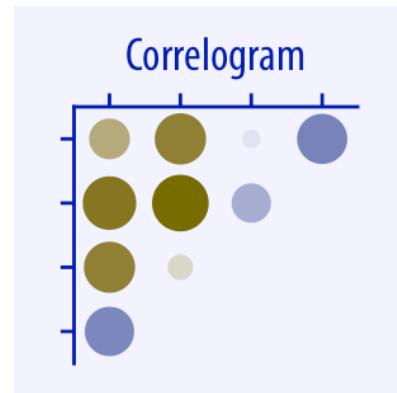
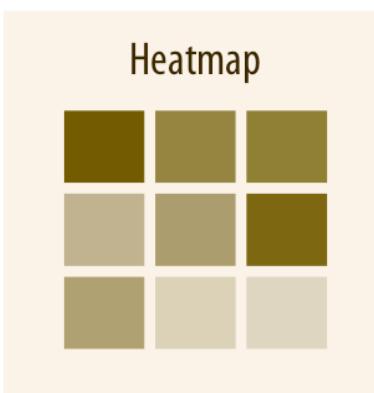
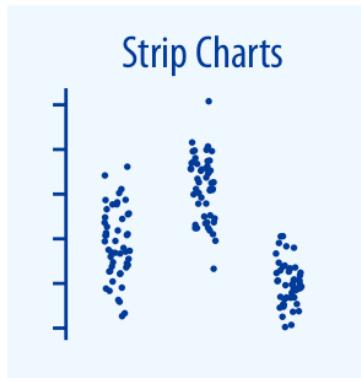
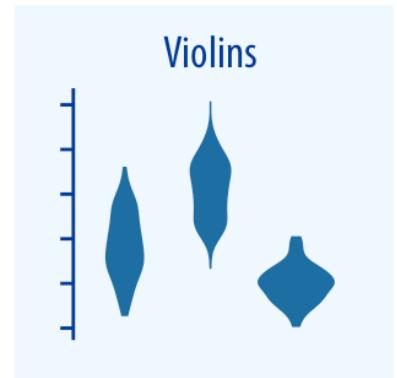
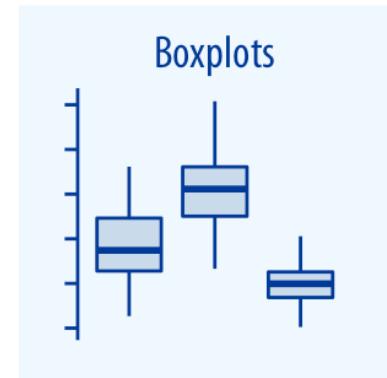
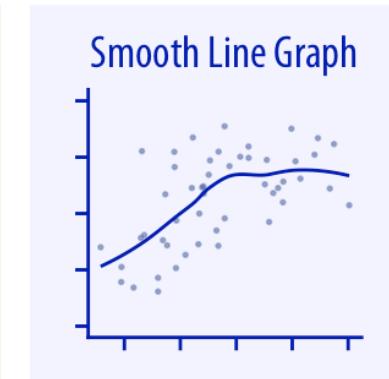
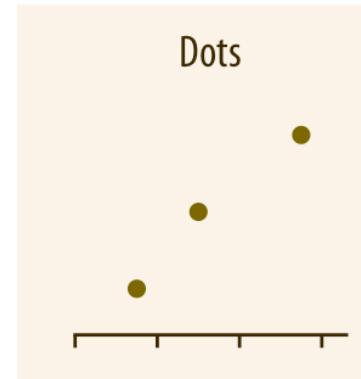
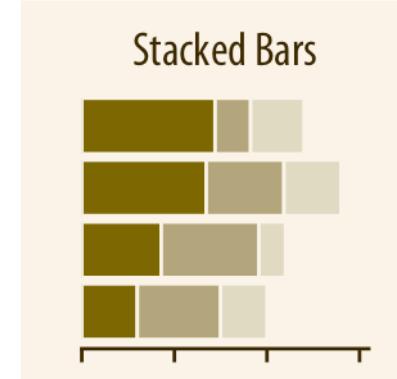
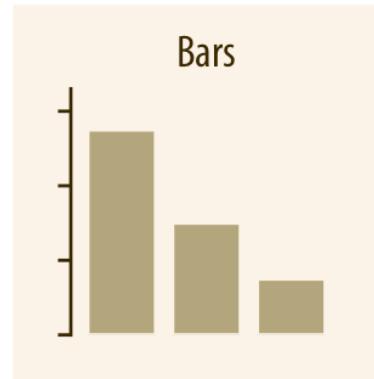
X-axis: Impression of longitudinal study, but everyone was asked same day. So, showing a “change over time” isn’t really appropriate.

World Values Survey: 1-10 scale, not Y/N!
Here **Yes == 10 == “Absolutely Important”**,
9-8 were ranked as “No”!

The importance of data visualization



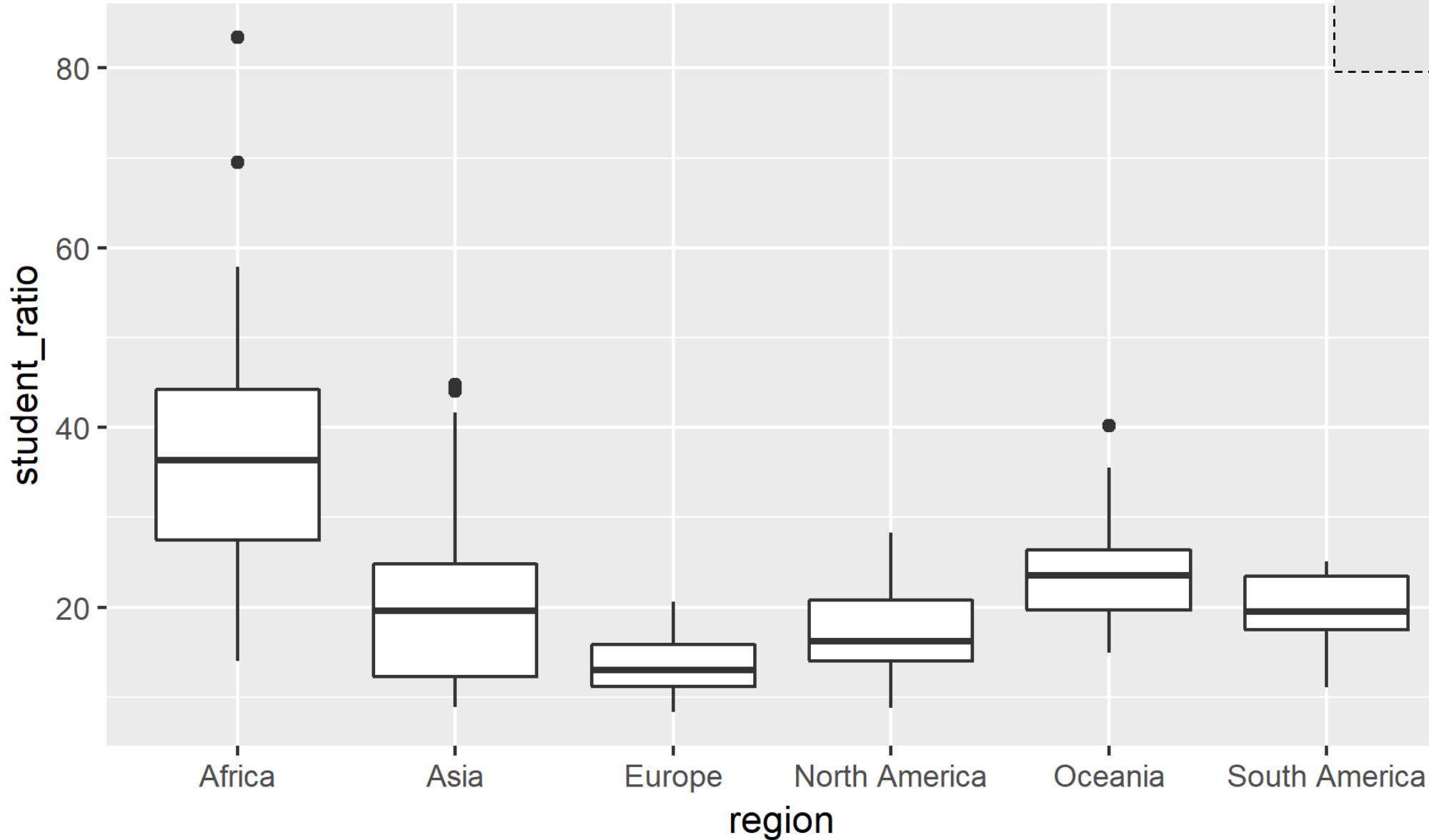
Maybe democracy will live to see another day, after all.



**Mix, match, and create
to your heart's content!**

Wilke, 2018

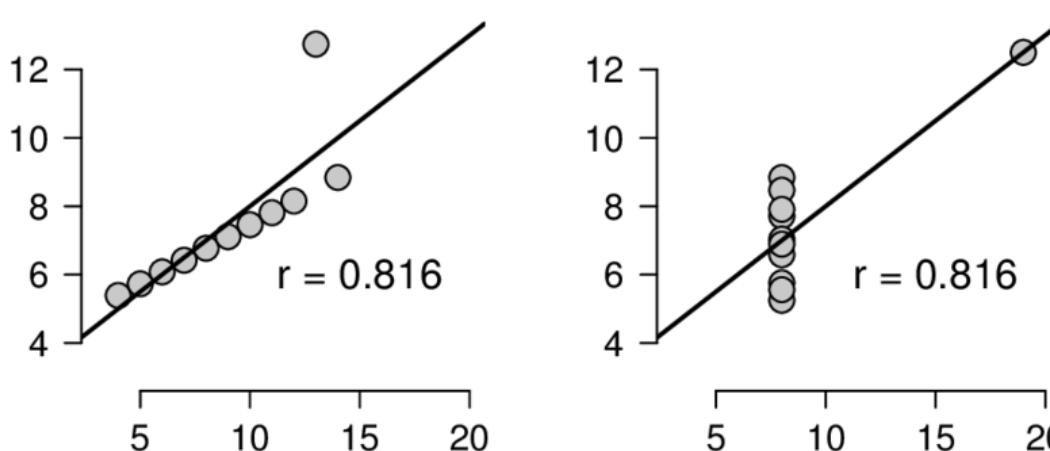
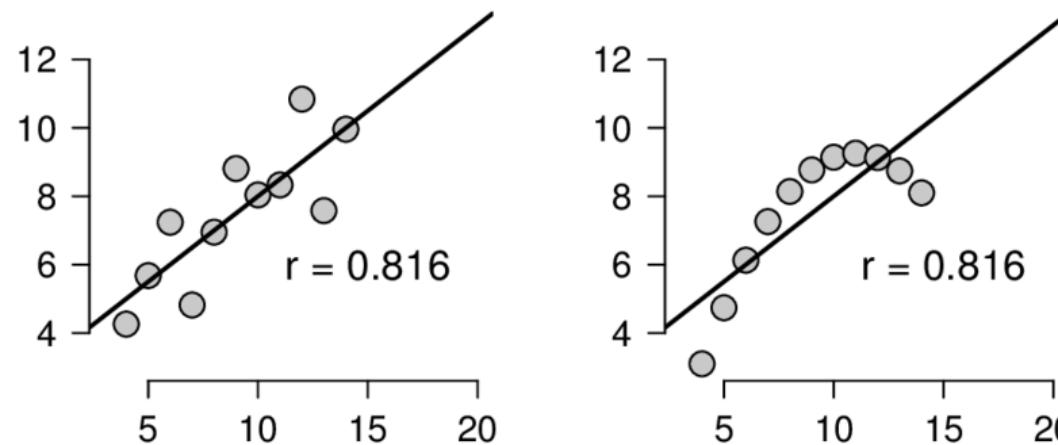
The Evolution of a ggplot



Data: UNESCO Institute for Statistics
Visualization by Cédric Scherer

Plotting: More than numbers

Anscombe's Quartet



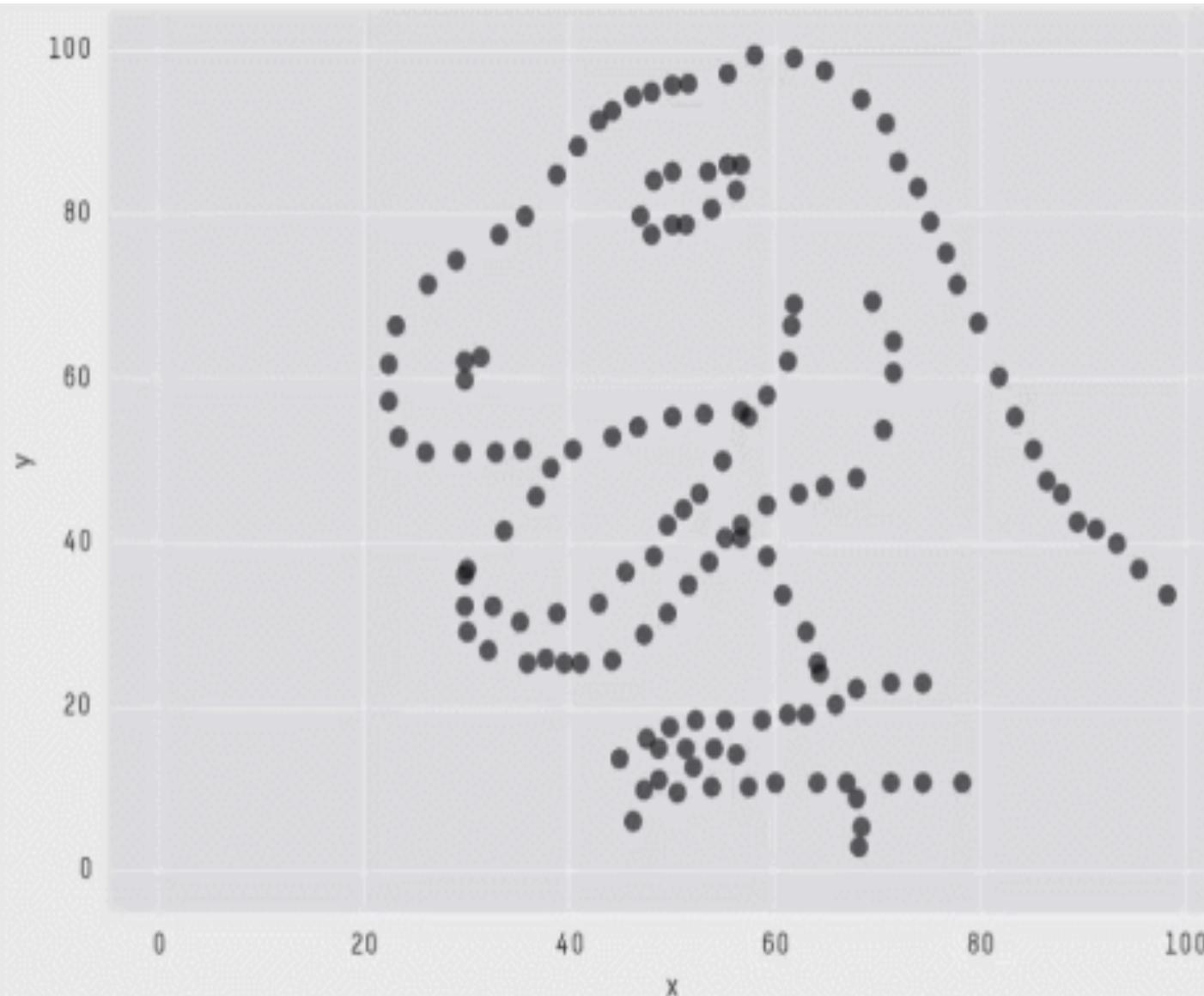
Working example: **anscombe** in R

All data sets here have identical means and correlations

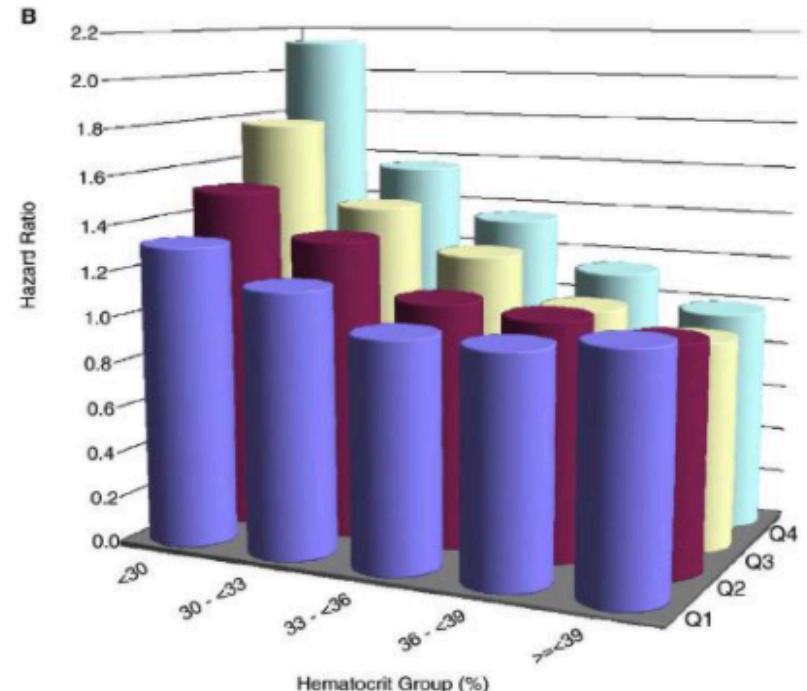
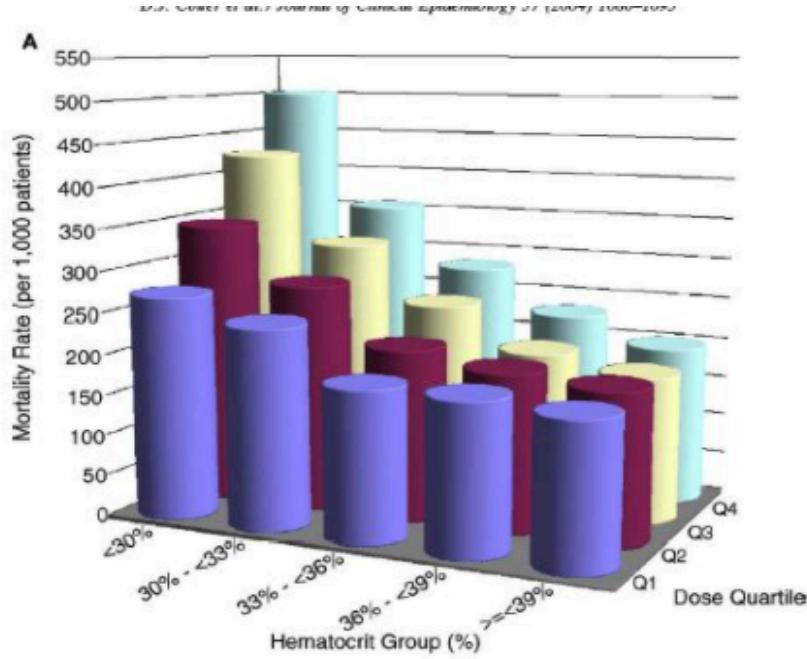
It's important to look consider both visual and mathematical descriptors of data

Datasaurus

Matejka and Fitzmaurice, 2017



X Mean: 54.2659224
Y Mean: 47.8313999
X SD : 16.7649829
Y SD : 26.9342120
Corr. : -0.0642526



How are differences being shown?

Aesthetics

Is it pretty?

Substantive

Is it correct?

Perceptual

Is it obvious?



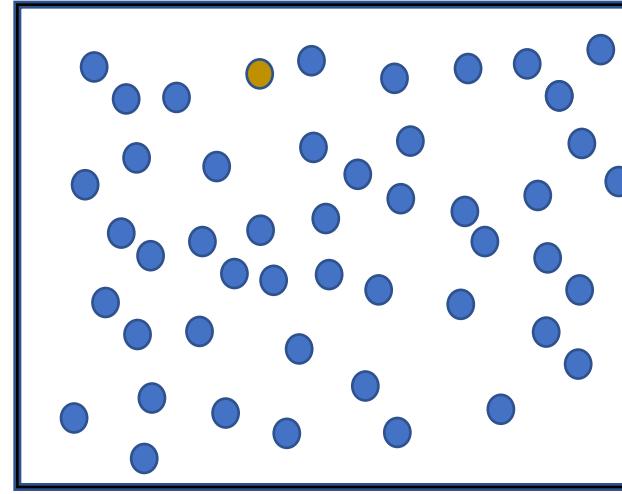
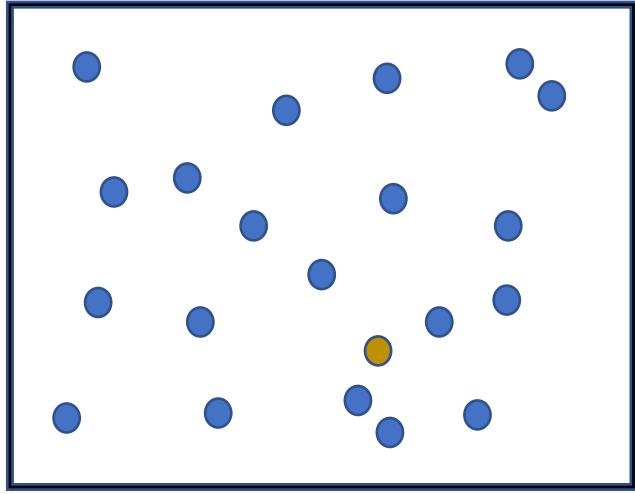
How would you improve this graph?

Aesthetics

Substantive

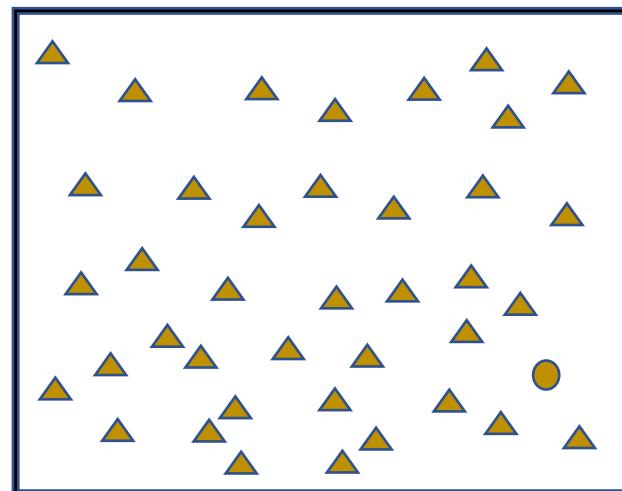
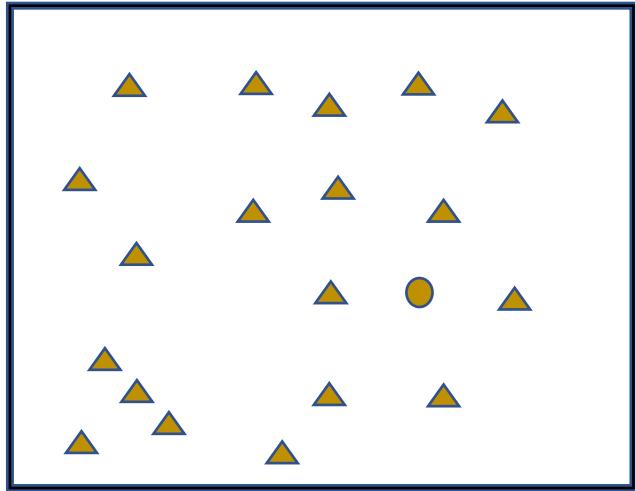
Perceptual

Information Channels

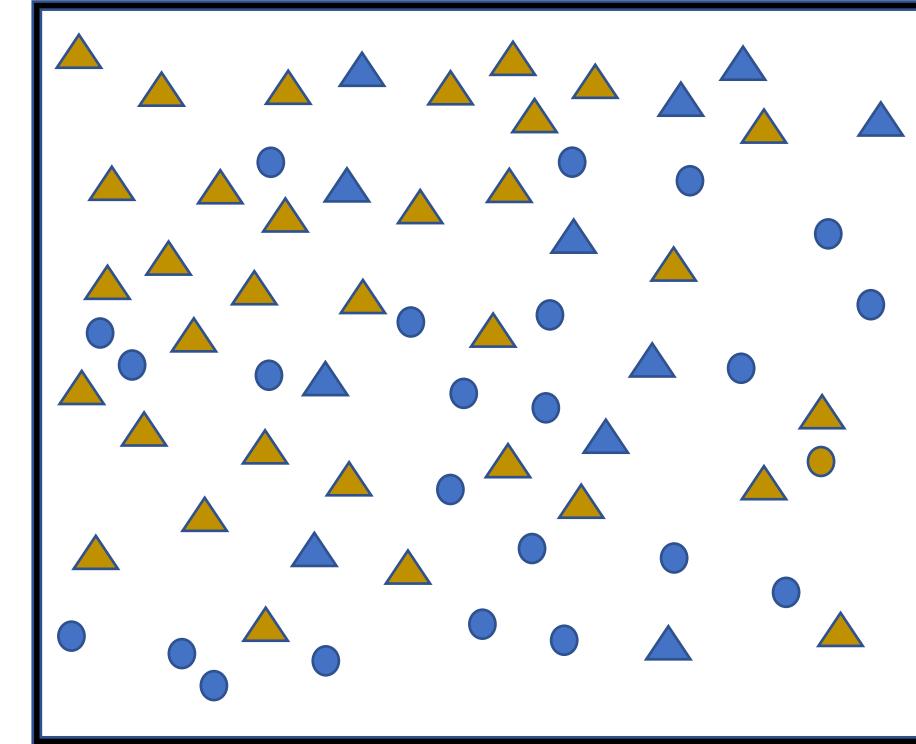


Task: Find .

Color



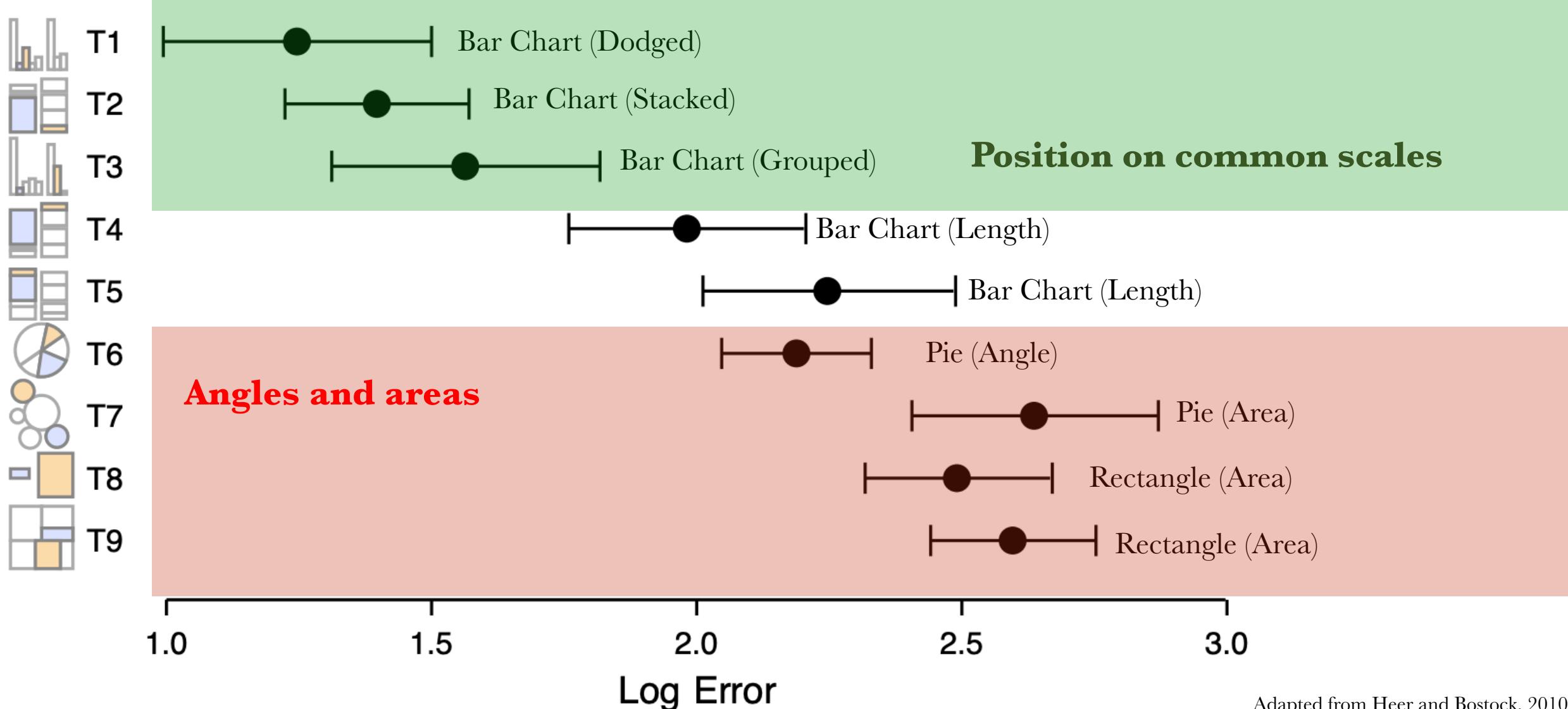
Shape



Both

Inspired by Healy 2019

Interpretation accuracy



Adapted from Heer and Bostock, 2010

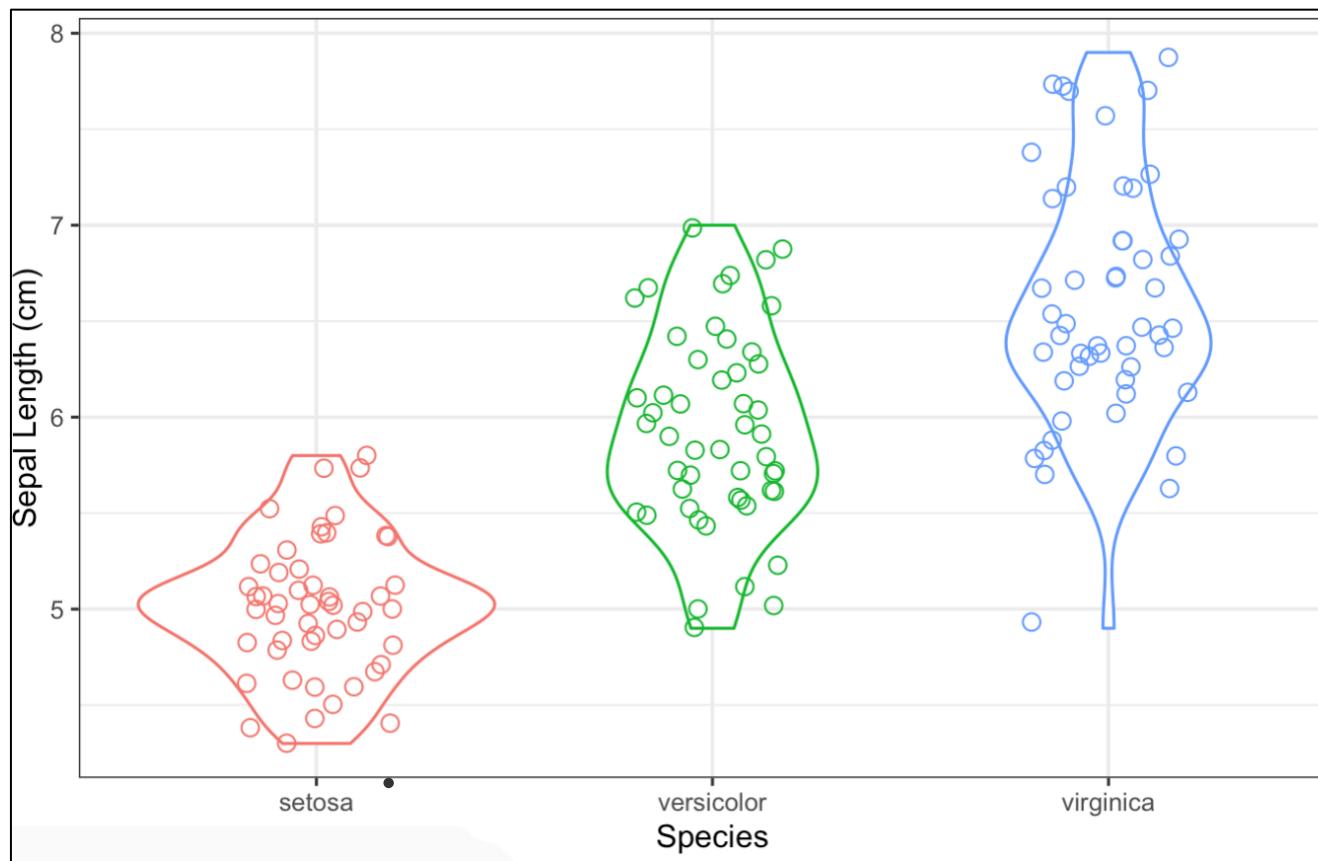


Welcome to the univeRSe

“Like all programming languages, **R** does exactly what you **tell it** to,
rather than exactly what you **want it** to.”

– *Data Visualization, Healy 2019*

Not great plot, Ex. 1



```
ggplot(data = iris,  
       aes(x = species, y = Sepal.Length)) +  
  geom_violin() +  
  geom_point() +  
  ylab("Sepal Length (cm)") +  
  theme_bw() +  
  theme(legend.position = "none")
```

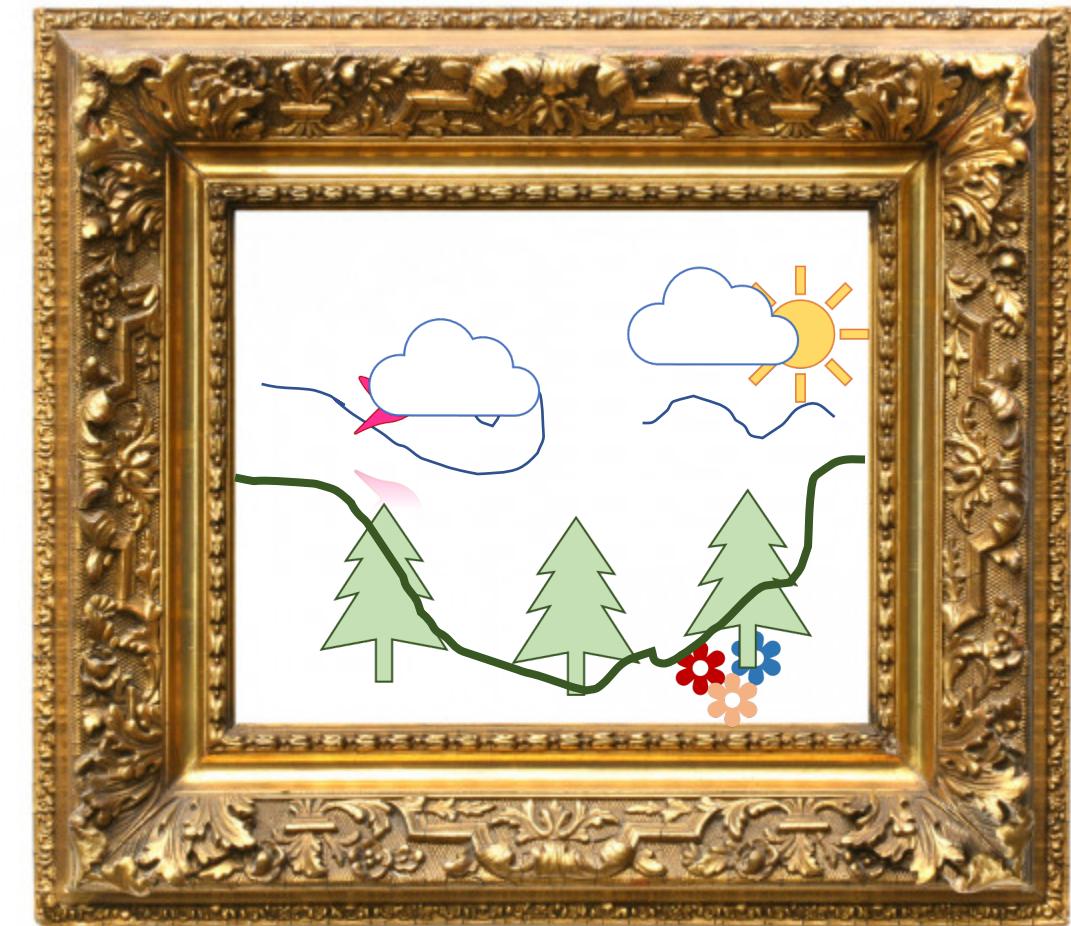
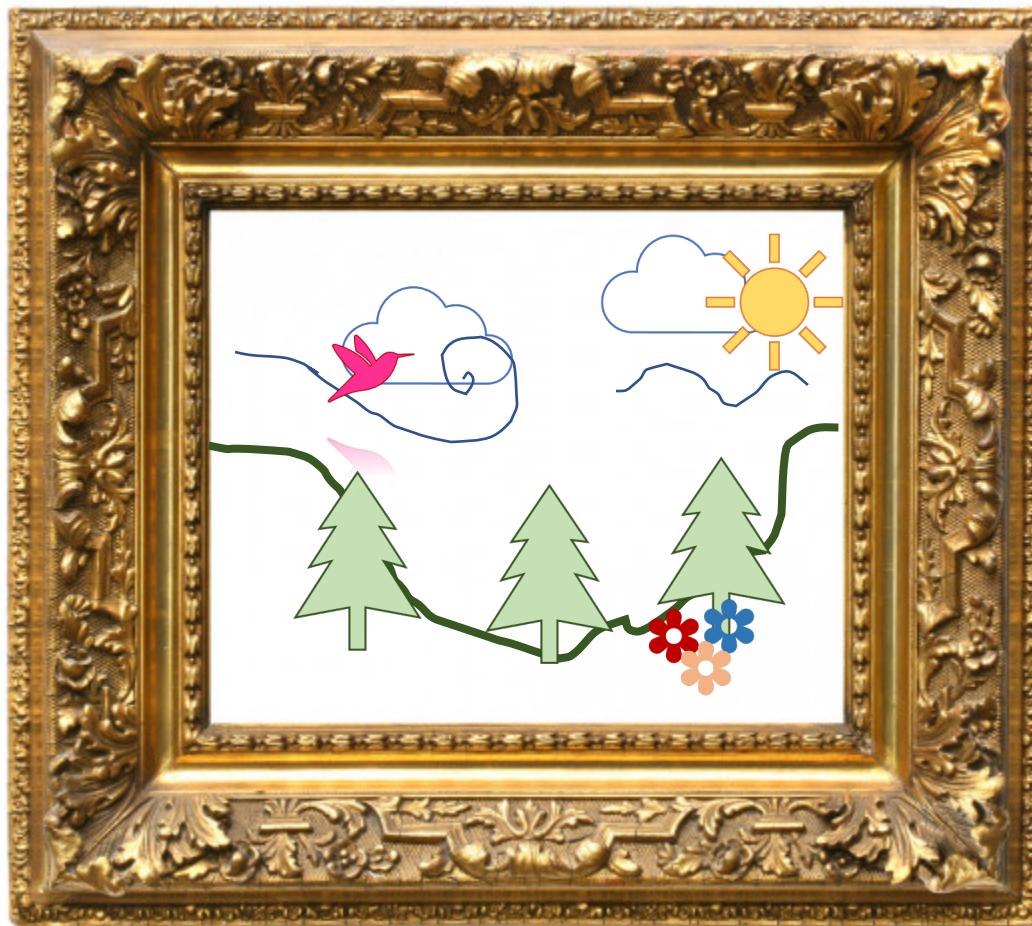
ggplot() Syntax

“...” = Character String, if **not** defined, non-code labels
ex. Name your axes

aes() = aesthetics
Assign **your data** to plot geography (x, y, size, color, shape).

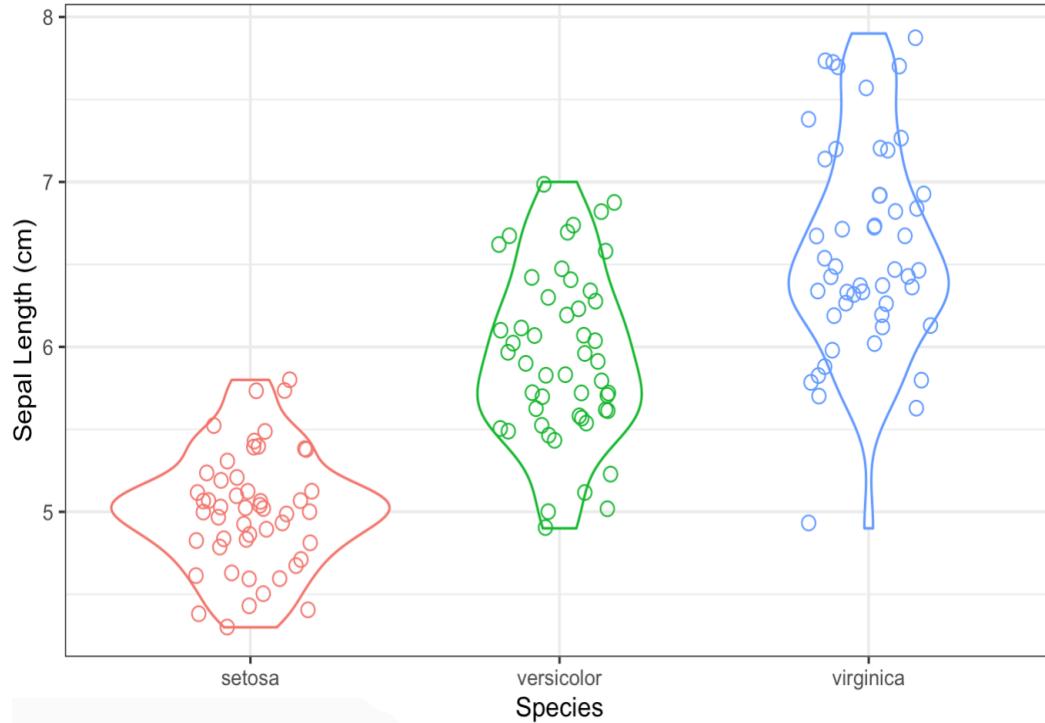
“...” = String **Defined**, performs specific task based on function
ex. Function and package specific.

ggplot() is a blank canvas



The order of your geometry matters in plot code!

- geom_violin() paints over geom_point()
- theme_bw() paints over theme(), keeping legend



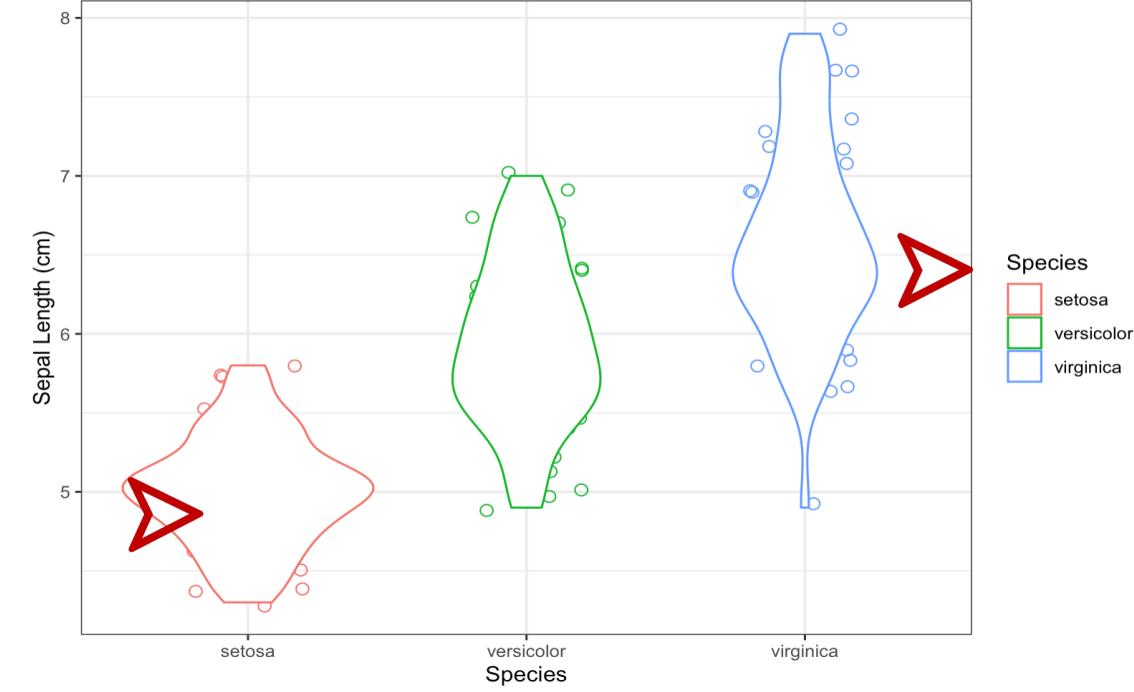
```
ggplot(data = iris,
       aes(x = species, y = Sepal.Length))+
  geom_violin() + ←  

  geom_point() + →  

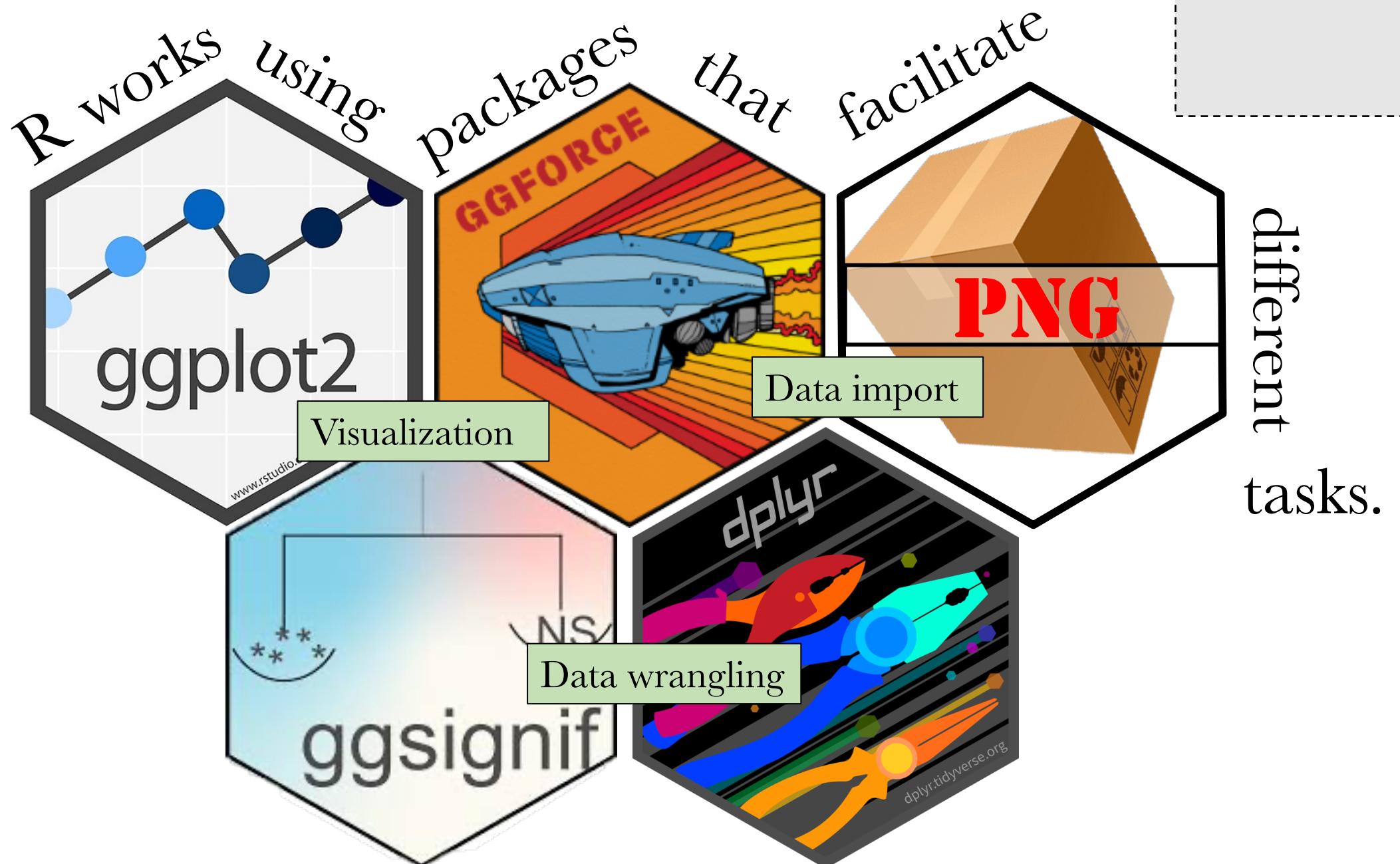
  ylab("Sepal Length (cm)") +  

  theme_bw() + ←  

  theme(legend.position = "none") →
```

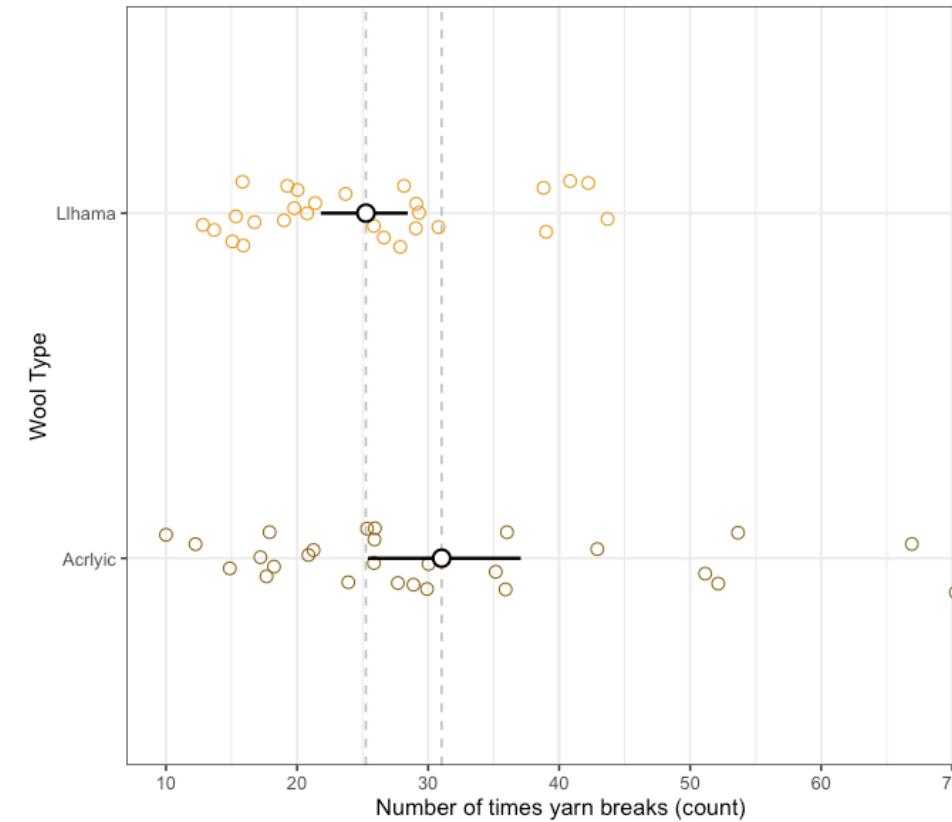
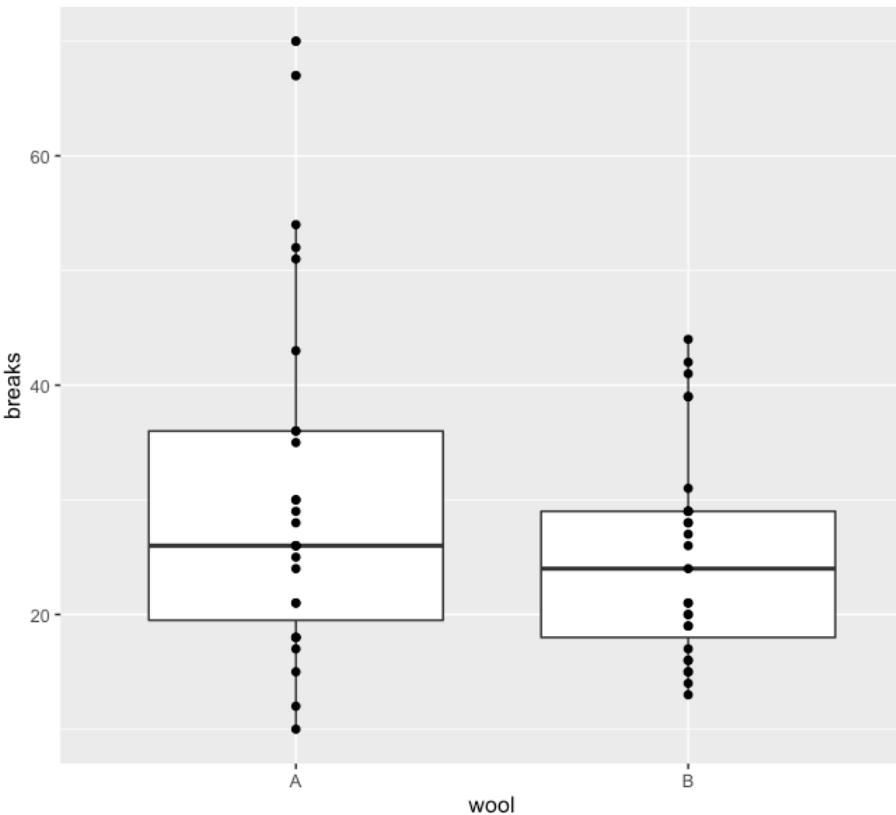


```
ggplot(data = iris,
       aes(x = species, y = Sepal.Length))+
  geom_point() +
  geom_violin() +
  ylab("Sepal Length (cm)") +
  theme(legend.position = "none") +
  theme_bw()
```



Part I: The Anatomy of `ggplot()`

Before



Boxplot

Color vs. Fill

Shapes, Sizes

Point, Jitter

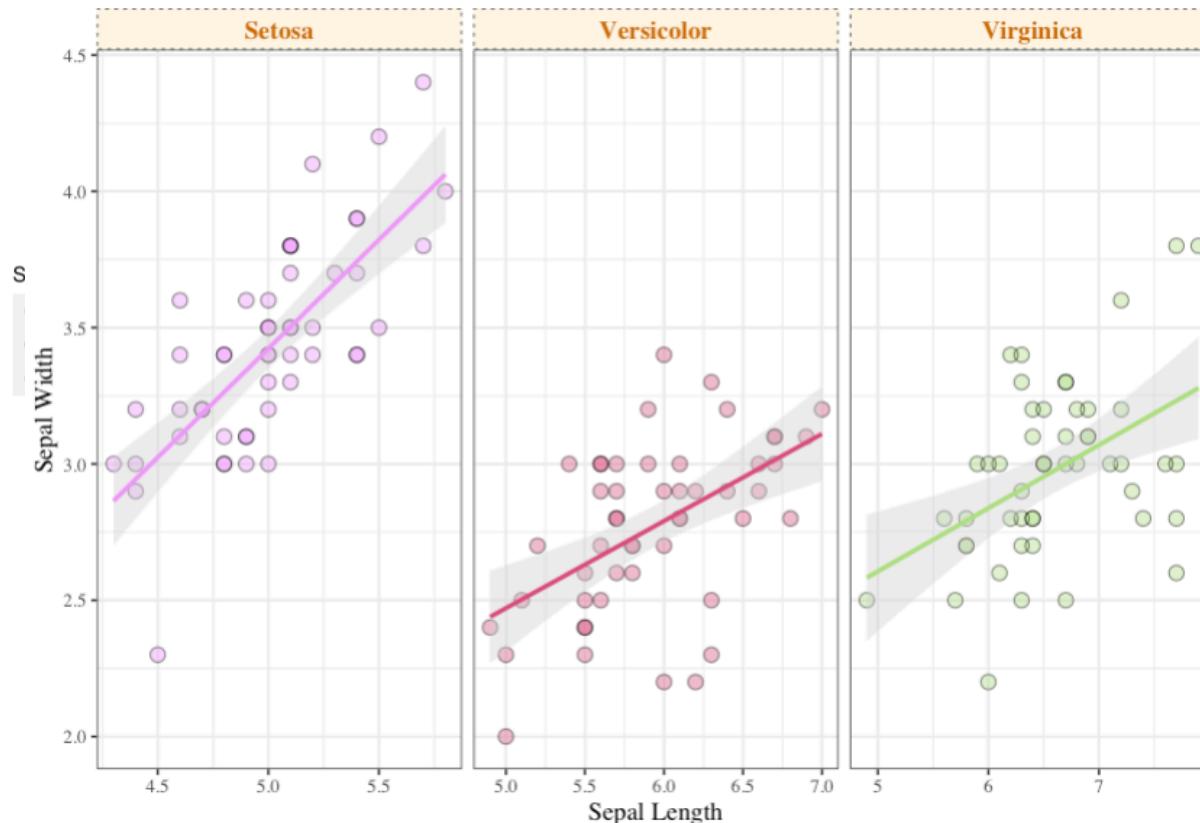
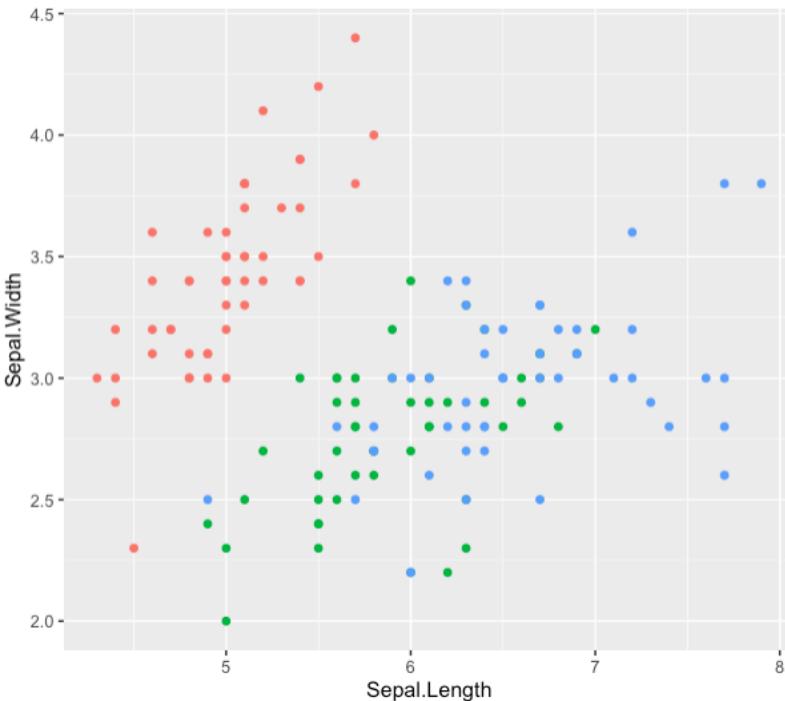
Custom color

Color brewer

Stat summary

Part II: Advanced customization

Before



Regressions

Themes

Text

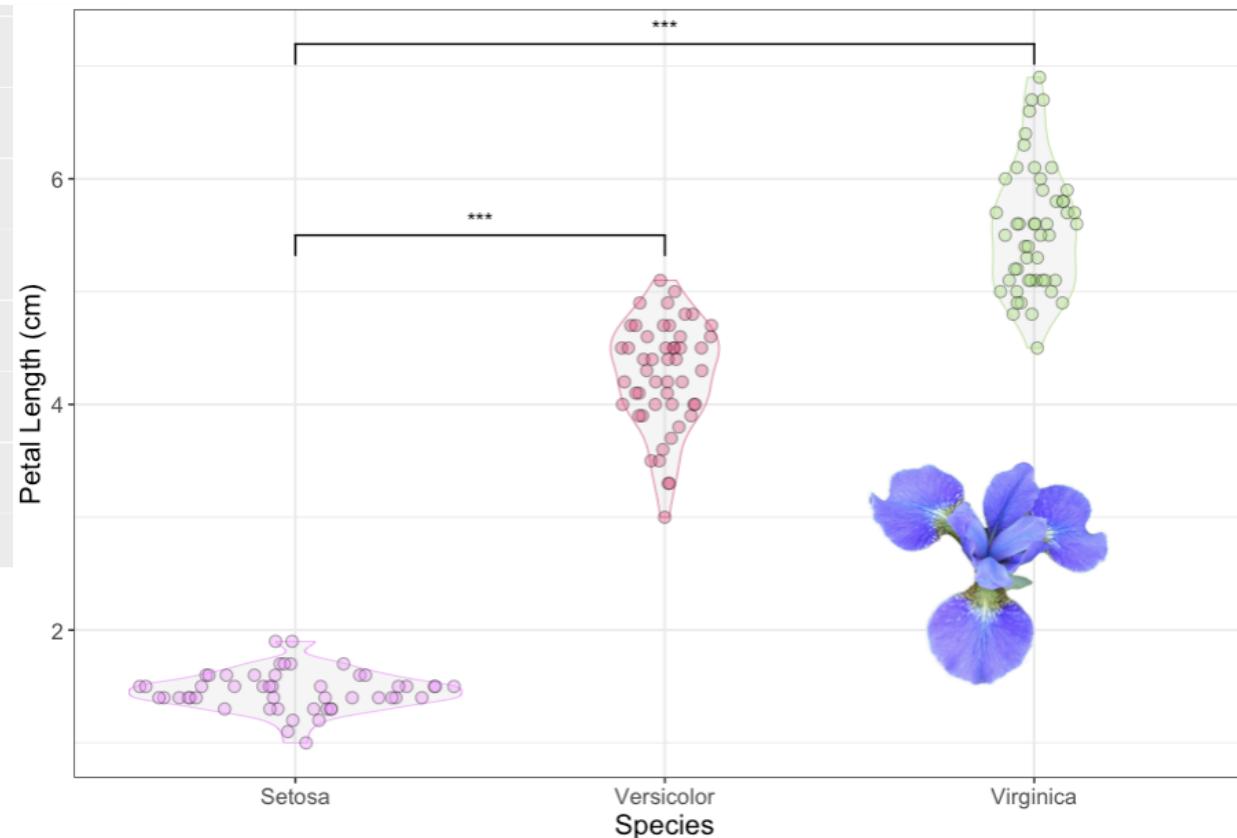
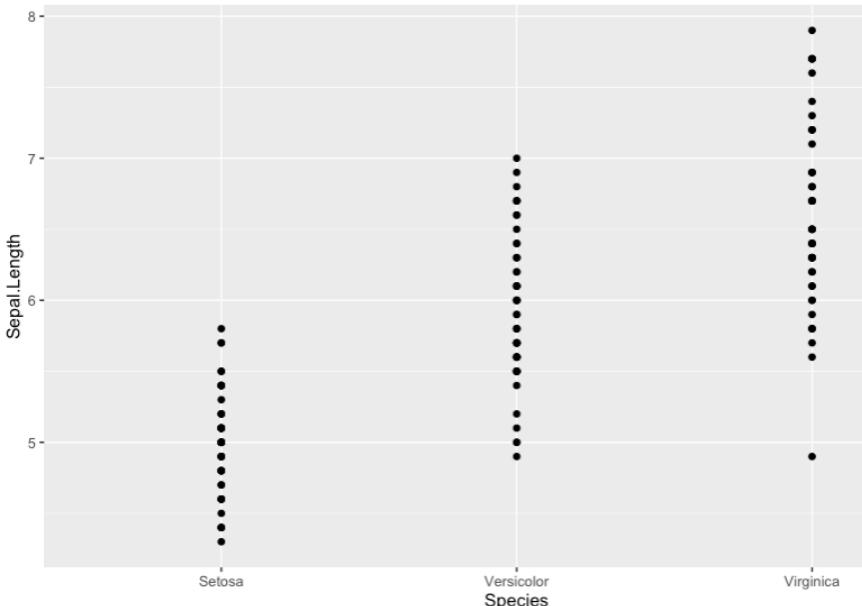
Background

Axis

Faceting

Part III: MasteR visualization

Before



Violin plots

Sina plots

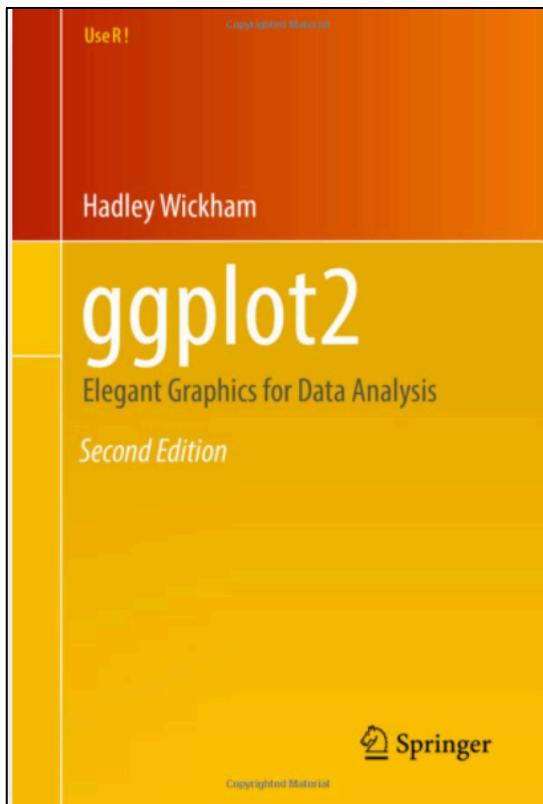
Sig. stars

Images

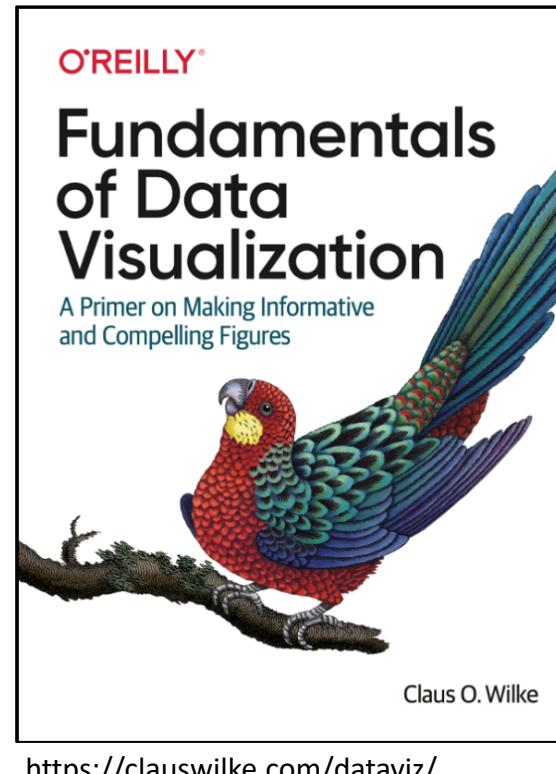
A GGPLOT2 TUTORIAL FOR BEAUTIFUL PLOTTING IN R

POSTED BY CÉDRIC ON MONDAY, AUGUST 5, 2019

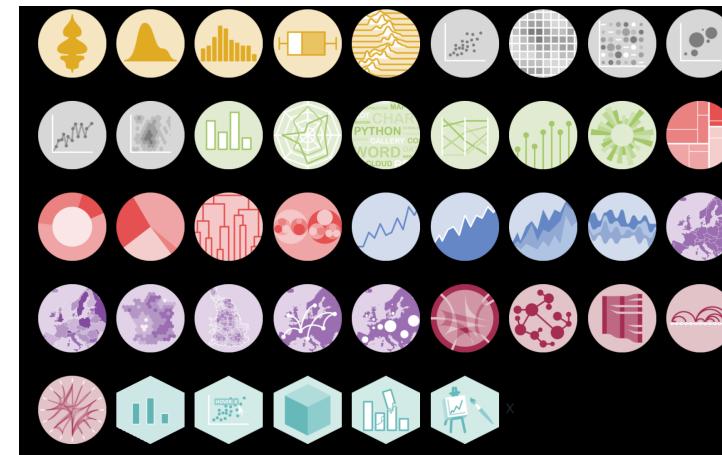
<https://www.cedricscherer.com/2019/08/05/a-ggplot2-tutorial-for-beautiful-plotting-in-r/#text>



<https://ggplot2-book.org/>



<https://clauswilke.com/dataviz/>

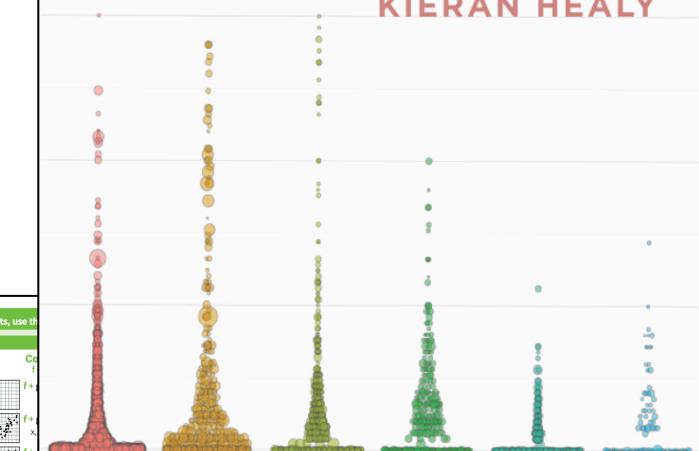


<https://www.r-graph-gallery.com/index.html>

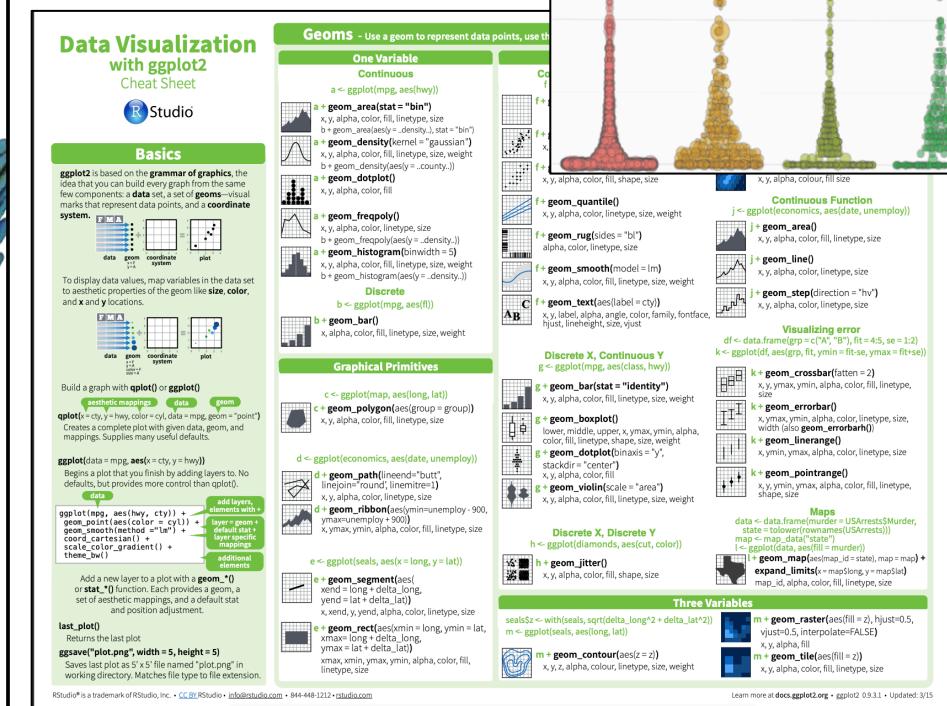
DATA VISUALIZATION

A PRACTICAL INTRODUCTION

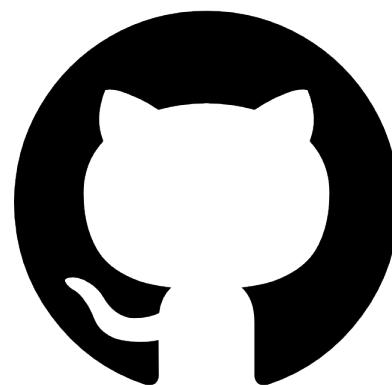
KIERAN HEALY



<https://socviz.co/index.html#preface>

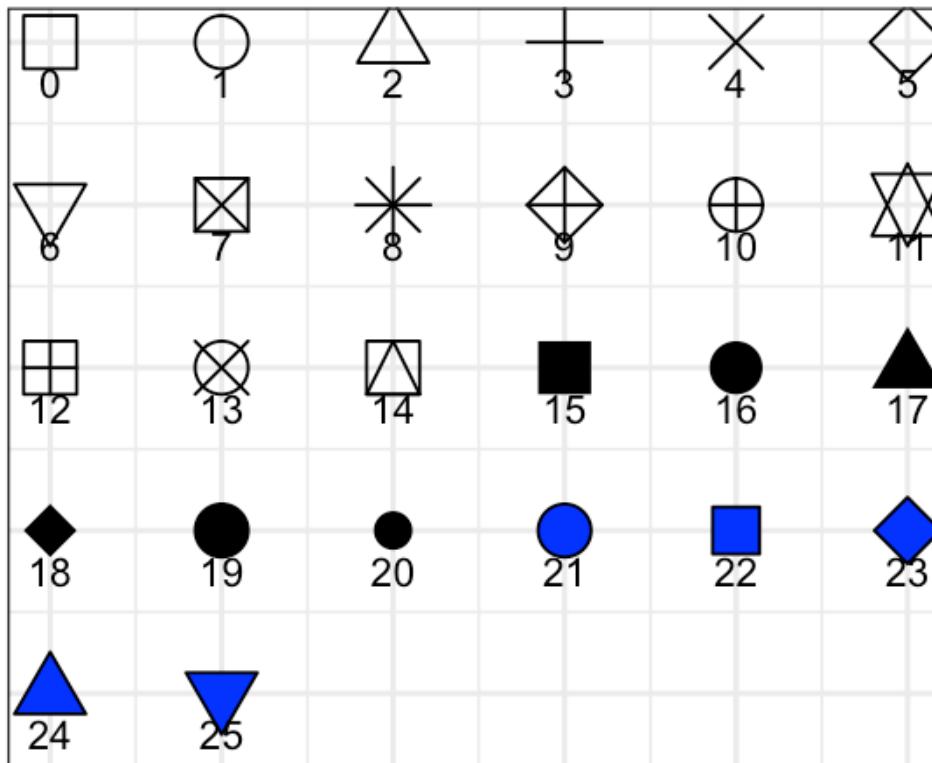


<https://rstudio.com/wp-content/uploads/2015/03/ggplot2-cheatsheet.pdf>



<https://github.com/Z3tt>

Point shapes available in R



Encycolorpedia

8d6502

Szukaj

Farba • Kolory w Internecie • Nazwy kolorów w HTML • Wykaz kolorów • Wsparcie

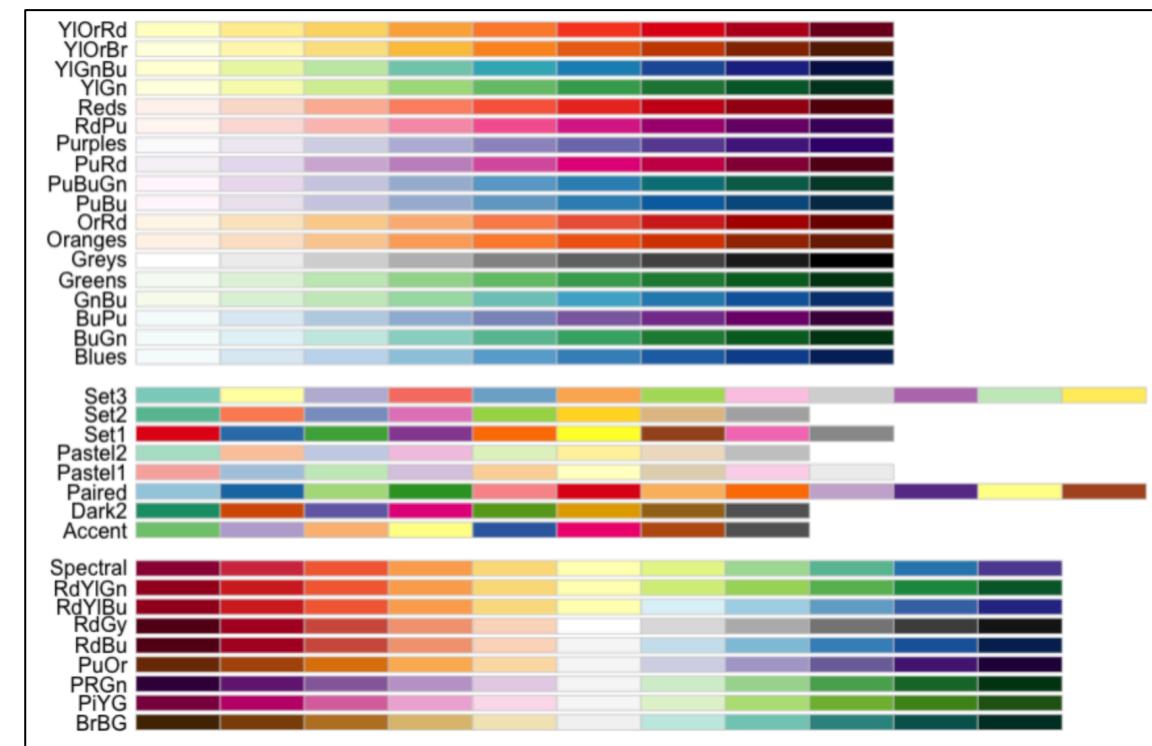


<https://encycolorpedia.pl/>

https://github.com/karthik/wes_anderson



Color Brewer Sets



Coffee break!

Don't you think there's a reason they don't call them "tea breaks"?

But that's none of my business.

The inferior
caffeinated beverage.

